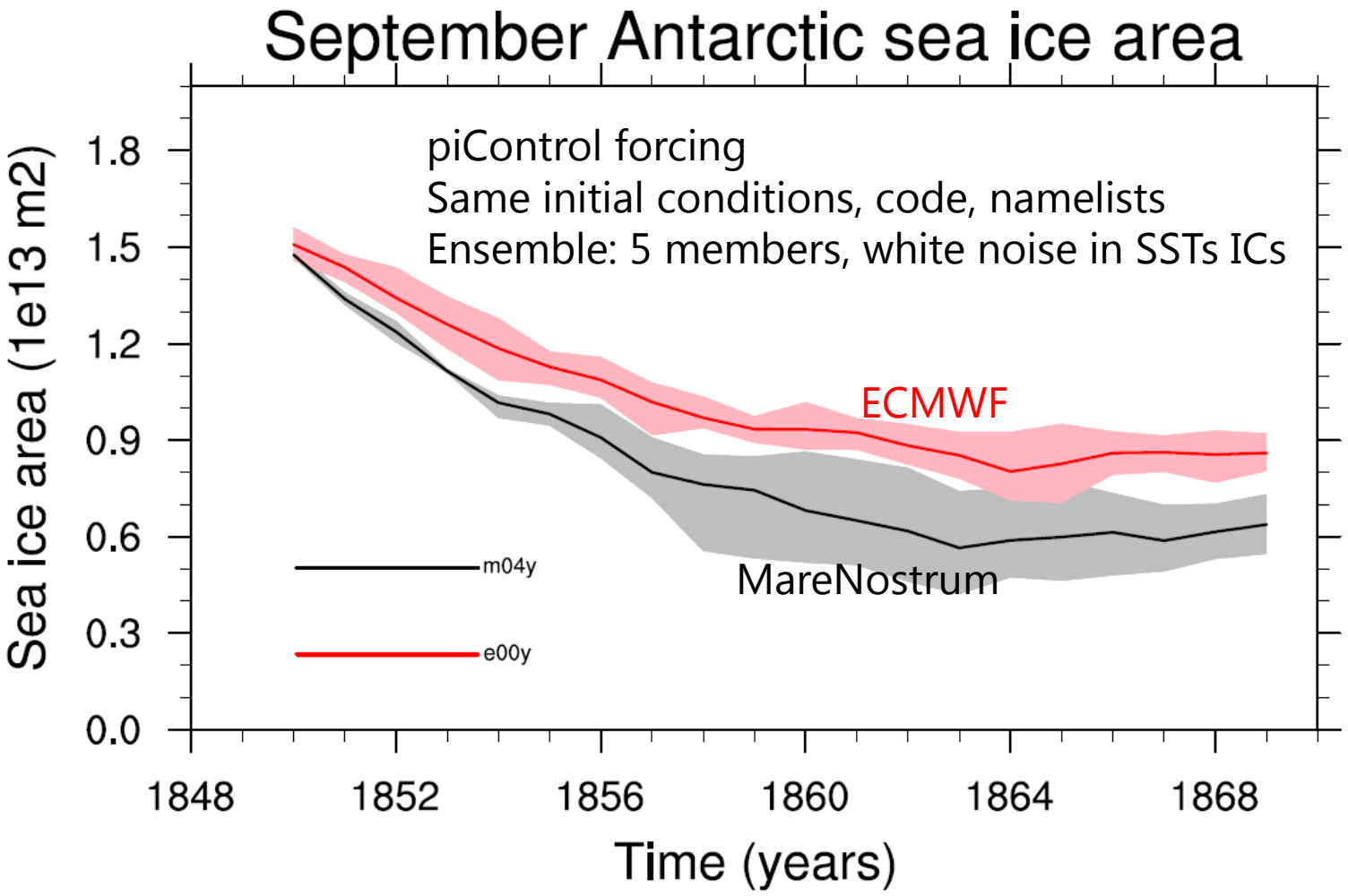
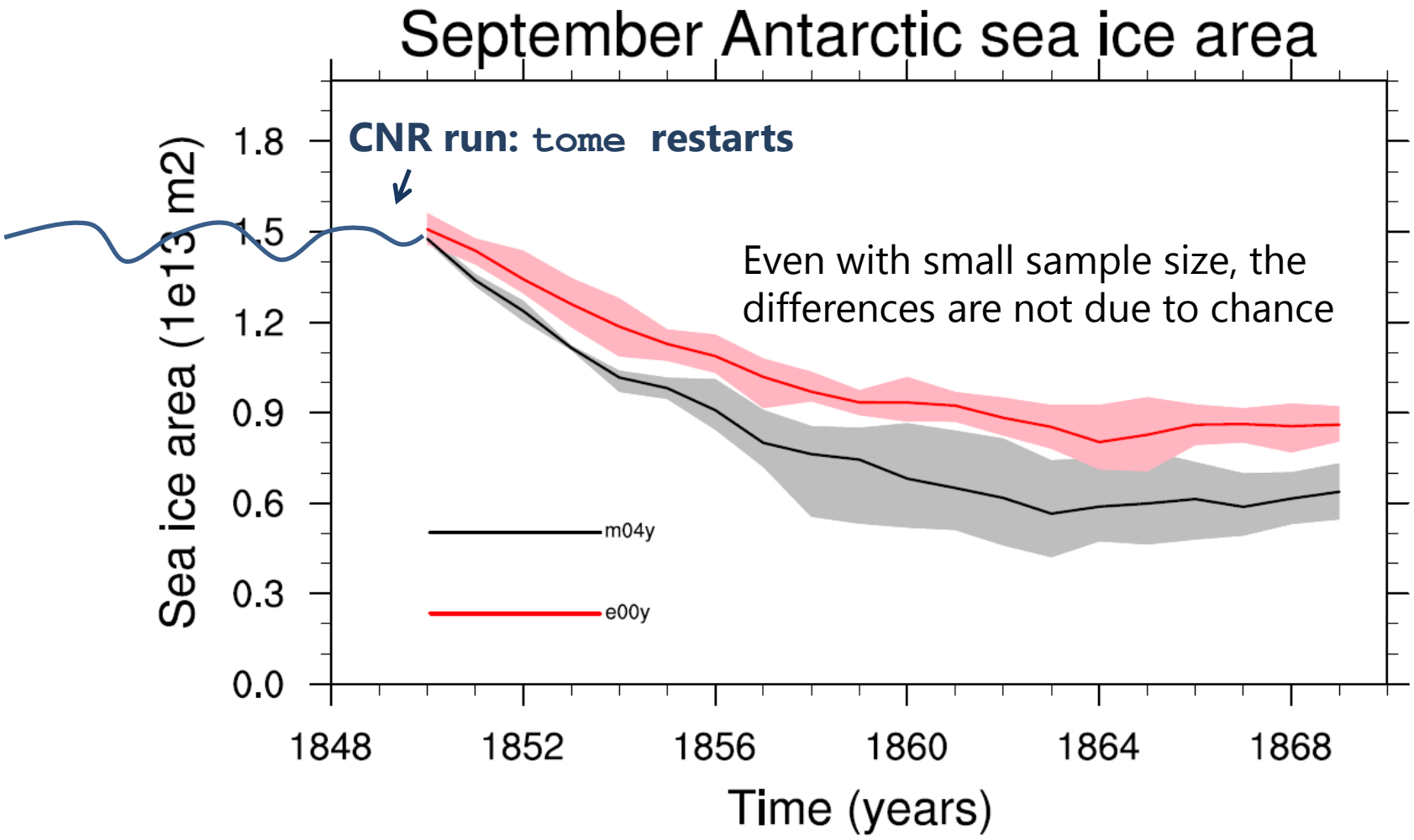
Reproducibility
2 Nov 2015

- Do different machines simulate different climates?
- Can we distribute expensive simulations on different platforms to reduce the burden?
- Do we underestimate uncertainty because of imprecise computers?



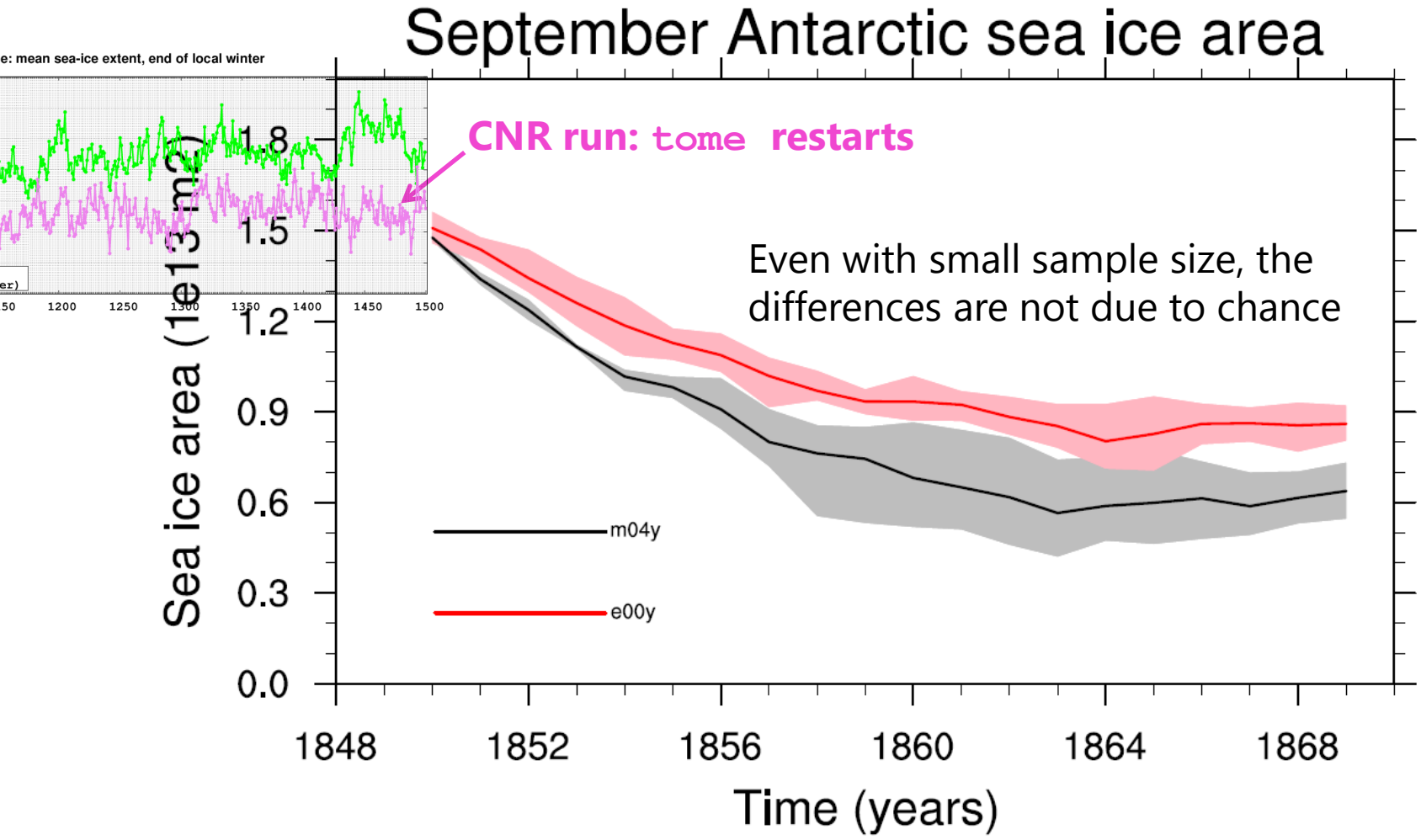
- Do different machines simulate different climates? **YES**
- Can we distribute expensive simulations on different platforms? **NO**
- Do we underestimate uncertainty because of imprecise computers? **YES**



BUT

Our setup isn't *exactly* perfect

- The model drifts despite starting from equilibrated restarts
- We use a different number and distribution of processors
- Key **`fp-model precise`** not enabled



BUT

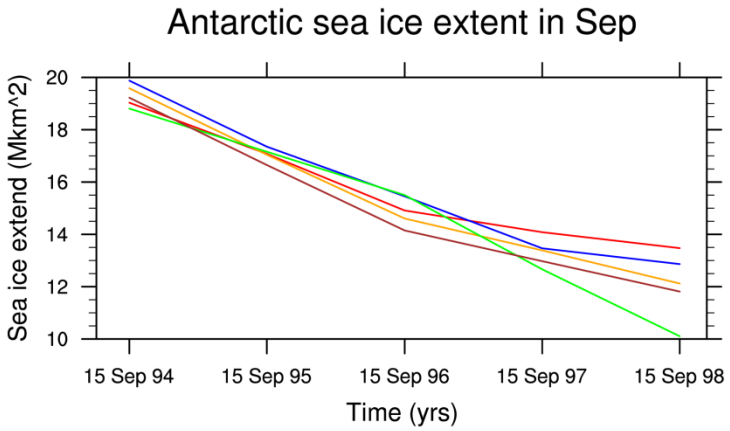
Our setup isn't *exactly* perfect

- **The model drifts despite starting from equilibrated restarts**
- We use a different number and distribution of processors
- Key **`fp-model precise`** not enabled

The model drifts despite starting from equilibrated restarts

Omar (a02f) and Eleftheria (i05m)'s simulations, also running with EC-Earth3.1, have the same type of issues

i05m



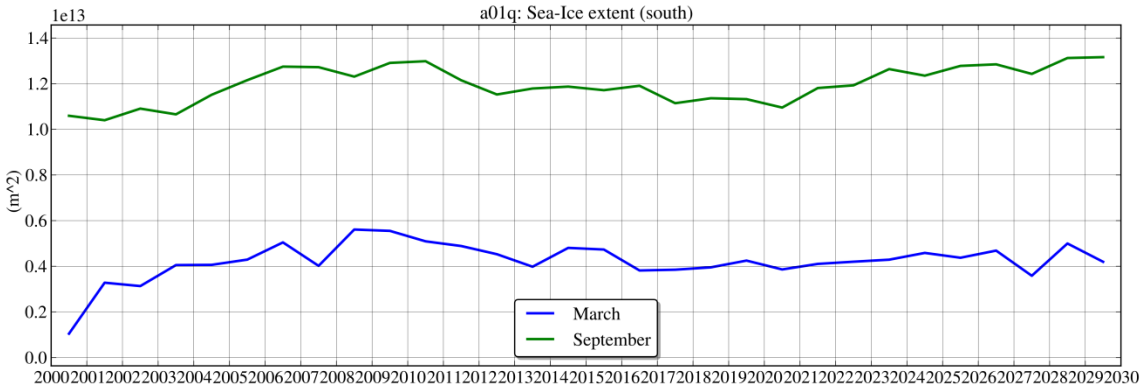
The model drifts despite starting from equilibrated restarts

- CNR and CFU/BSC use the same model version [r1692] but
  - IFS namelists differ: changes from Lauriane
  - NEMO namelists differ: changes from Klaus, Chloe, Uwe
  - Codes differ: patches from Klaus, Isabel AB, Uwe

Preliminary conclusion: **it does not make sense to exchange restarts with other groups even if we « checked out » the same version**. We all introduce (consciously or not) additional layers that make models different!

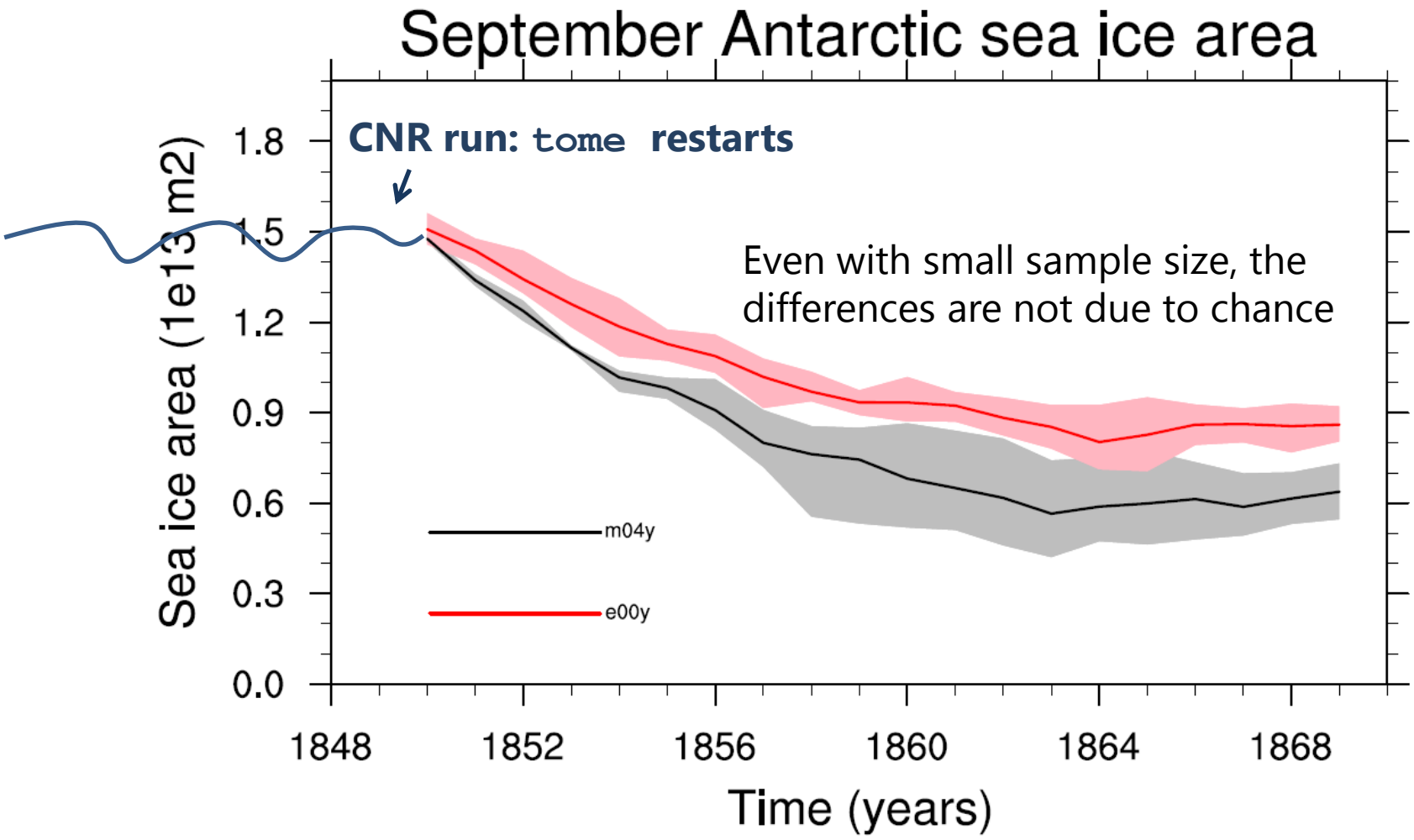
The model drifts despite starting from equilibrated restarts

- The loss of Antarctic sea ice is related to SSS issues
  - No loss if SSS restoring **and** E-(P+R) forced to zero



Here we face a dilemma.

- If flags for SSS restoring and E-(P+R) are not enabled, EC-Earth goes crazy in the Southern Ocean (and this likely affects the whole Southern Hemisphere)
- But it does not make sense to run decadal simulations with restoring



BUT

Our setup isn't *exactly* perfect

- The model drifts despite starting from equilibrated restarts
- **We use a different number and distribution of processors**
- **Key `fpprecise` not enabled**

Processors:

If their number and distribution really matters, that's a problem. Because we can't run with as many procs on Ithaca as on MN3.

→ we should level down all our simulations to the smallest machine

FP precise key

fp-model, fp

Controls the semantics of floating-point calculations.

Syntax

Linux OS and OS X:	`-fp-model` *keyword*
Windows OS:	`/fp:`*keyword*

Arguments

keyword	Specifies the semantics to be used. Possible values are:
`precise`	Disables optimizations that are not value-safe on floating-point data and rounds intermediate results to source-defined precision.
`fast[=1|2]`	Enables more aggressive optimizations on floating-point data.
`strict`	Enables precise and except, disables contractions, and enables the property that allows modification of the floating-point environment.

Conclusions: problems are much deeper than the simple question of reproducibility

EC-Earth 3.1 has a huge bias
Can we trust its Southern Hemisphere climate?

EC-Earth 3.1 is *not* reproductible under current standards
Even with autosubmit, and a dedicated experiment!

Is EC-Earth 3.1 reproductible under very strict standards?
New experiments are to be started. But what do we do with conservation keys?

And all this should be repeated with EC-Earth 3.2!